

**Optimize Data Science
Workflows at Scale**

Optimize Data Science Workflows at Scale

A Habu data clean room adds critical value to data science workflows, particularly at scale. We unpack all the ways Habu supercharges your data-driven initiatives.

For enterprises that are — or aim to be — data-driven, data science teams play a crucial role: they enable organizations to extract actionable insights from vast amounts of data. These insights can then be used to optimize business processes, enhance customer experiences, identify new opportunities, and ultimately gain a competitive advantage in the marketplace. It's via data science that enterprises can truly unlock the full potential of their data assets, fostering innovation and powering business success in an increasingly data-centric world.

A modern data clean room is a critical tool for enterprises and their data science teams to power data-centric objectives and put their data to work. The most advanced data clean room software, like Habu's, is a powerful tool for streamlining and optimizing data science workflows, especially when working with large datasets and multiple partners.

At a high level, Habu data clean room provides your data science teams with three key capabilities that accelerate productivity and drive value:

- 1 Enrich first-party data** via second-party data from partners to build high-quality training datasets without needing any data engineering skills.
- 2 Easily spin up clean rooms** for your partners — and execute models directly inside these clean rooms for a variety of use cases.
- 3 Leverage unprecedented simplicity** to give business users access to data and insights while slashing support requirements.



Unpacking Habu's Advantages

Let's dig a little deeper into how Habu's core capabilities make it an essential tool for optimizing data science workflows in an era of decentralized data and privacy-first mandates. We'll begin with the best-in-class features that deliver seamless data collaboration:

Engage in One or More Clean Rooms at Scale

For maximum flexibility, data science teams need to be able to work fluidly with multiple data partners. While you could manually build out access to multiple partner clean rooms, that's a slow and tedious process. Habu's automation and orchestration capabilities let you easily work with many clean rooms at scale — and add more as needed. Think of Habu not as infrastructure, but as an enabling software layer (control plane) on top of a partner's enterprise or walled garden clean room (data plane). Habu makes it possible for data scientists to analyze data across numerous sources, enabling better insights and decision-making.

Access Data You Would Not Otherwise Get

Controlled, secure data sharing in a clean room environment is a requirement for many data owners who are unwilling to share their data through other, less secure means. Providing compliance with privacy regulations and data protection standards, Habu enables organizations to access valuable data that they would not be able to obtain without a clean room solution.

Leverage Cloud Interoperability

Habu's cloud interoperability allows data scientists to work with a broad range of cloud providers without being constrained by technical complexity. This flexibility enables organizations to choose the best provider for their needs and facilitates collaboration with partners who may be using different cloud platforms.

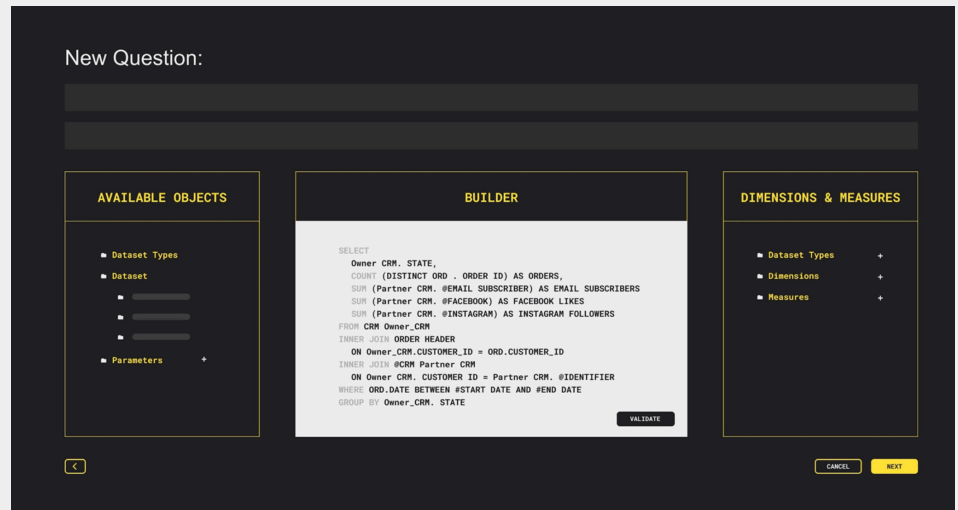
Enjoy Unrivaled Simplicity

Free your teams to accelerate. Habu significantly reduces the need for IT support in a clean room environment. Its user-friendly interface allows data scientists to focus on their work without getting bogged down in technical details. And Habu's simplicity also makes it easy for lines of business to operate, ensuring that data-driven insights are accessible across the organization.



Special Features Built For Data Science

Habu's industry-leading capabilities are perfect for the needs of data science teams. But Habu's solution goes further — with three advanced features designed specifically to accelerate data science initiatives:



Habu Question Builder

Habu CleanCompute: Go Beyond Simple Insights

Enrich your data, or your partner's data, with machine learning models and advanced computing. Habu CleanCompute provides a secure environment in which to run virtually any containerized code. This feature allows data scientists to leverage state-of-the-art packages and libraries for machine learning and deep learning use cases. These include image recognition, fraud detection, and other advanced applications where you need to extract deeper insights from data.

Habu CleanML: Power Your Machine Learning Workloads

CleanML is specifically built for machine learning tasks. Unlike CleanCompute, CleanML is not containerized, making it simpler and more efficient for running ML workloads. Use the data coming out of Habu data clean room to train models, and then immediately output a trained model. Or bring in a trained model to draw inferences. You can even do both together!

Habu Question Builder: Distribute Insights at Scale

Write your own queries and templatzite them for re-use across datasets. Use the same question across multiple clean rooms, organizations, and accounts to avoid manually repeating your work. Question Builder uses a natural language framework, making it easy for business users to find the reports they want and trigger their own runs. They can even insert their own runtime parameters, enabling a high degree of customization.



Many organizations do not have an in-house data science team. Habu data clean room assists these customers with easy setup and out-of-the-box queries that enable them to quickly get started mining data for insights. With a user-friendly interface and intuitive tools, even businesses without dedicated data science teams can leverage the power of Habu to drive data-driven decision-making.



Habu Data Clean Room Is Unlike Anything Else

Compared to other clean room vendors, Habu software stands out in important ways.

Data Science Teams

Low-code for technical users

Create & distribute analytics and outcomes powered by 1P and partner data – at scale

Flexible

Configurable

Build your own

Open ended



Business Teams

No-code for business users

Receive valuable insights and outcomes at the push of a button and leverage the results anywhere

Pre-packaged queries

Intuitive user interface

Plug & Play

Automated

With Habu, there is no data movement — you connect to data where it lives and go. Besides being a requirement for many data owners, no data movement reduces latency and the risk of potential data leakage. Habu's cloud interoperability enables data partners to easily connect to you, facilitating seamless data access and collaboration.

Habu is also API-driven, making it simple to integrate with your existing workflows and systems. Its built-in, enterprise-grade BI tool allows users to visualize and analyze data without needing additional software.

Easily scale not just the environment (your portfolio of clean rooms), but the questions, models, and training as well — all to make it easier for business users to consume. This scalability ensures that as your organization's data needs grow, Habu can accommodate those needs without significant overhead or complexity.



Seamless Data Collaboration to Power Data Science

Habu data clean room is a powerful, scalable solution that transforms the way organizations approach data science workflows. Its unique features and capabilities enable data scientists to dramatically simplify the process of accessing and analyzing data from multiple sources. With a secure, easy-to-use environment, Habu empowers organizations to optimize data science workflows at scale and derive valuable insights from their data.

How does this play out in the real world? To take just one example, the data science team of a global CPG company leveraged Habu's data clean room software to embrace agile experimentation into new use cases, technologies, and partners. First, they leveraged Habu's interoperability to easily tap into the large walled garden clean rooms — Facebook, Amazon, and Google — and dramatically expand their access to data. They then connected to retail partners around the world in order to enrich their first-party data, drive ad performance, and improve marketing ROI. Finally, the team launched a large cross-platform optimization project with Habu that leverages enriched 1st-party data and partnerships to improve supply chain, product manufacturing, and more.

The result? The CPG company, which operates diverse brands in markets around the world, dramatically improved access to data and measurement of marketing efficiency while emerging as a truly innovative global enterprise.

To see the impact of Habu data clean room on a diverse range of industry-leading brands, explore our case studies.

To experience the power of Habu
for yourself, simply
[request a demo.](#)



Habu enables organizations to master decentralized data, uncovering insights that fuel better business decisions via data collaboration that's smart, safe, scalable, and simple. With flexible, multi-cloud deployment, high levels of automation, and an intuitive interface, Habu Data Clean Room empowers data teams. Easily connect your data from any source and start accelerating growth today.